

Computer lab 3

Instructions

- Create a report to the lab solutions in PDF.
- Be concise and do not include unnecessary printouts and figures produced by the software and not required in the assignments.
- **Include all your codes as an appendix into your report.**
- A typical lab report should 2-4 pages of text plus some amount of figures plus appendix with codes.
- The lab report should be submitted via LISAM before the deadline.

Assignment 1: Cluster sampling

An opinion pool is assumed to be performed in several locations of Sweden by sending interviewers to this location. Of course, it is unreasonable from the financial point of view to visit each city. Instead, a decision was done to use random sampling without replacement with the probabilities proportional to the number of inhabitants of the city to select 20 cities. Explore the file **population.xls**. Note that names in bold are counties, not cities.

1. Import necessary information to R.
2. Use uniform random number generator to create a function that selects 1 city from the whole list by the probability scheme offered above (do not use standard sampling functions present in R).
3. Use the function you have created in step 2 as follows:
 - a. Apply it to the list of all cities and select one city
 - b. Remove this city from the list
 - c. Apply this function again to the updated list of the cities
 - d. Remove this city from the list
 - e. .. and so on until you get exactly 20 cities.
4. Run the program. Which cities were selected? What can you say about the size of the selected cities?
5. Plot one histogram showing the size of all cities of the country. Plot another histogram showing the size of the 20 selected cities. Conclusions?

Assignment 2: Different distributions

The double exponential (Laplace) distribution is given by formula:

$$DE(\mu, \alpha) = \frac{\alpha}{2} \exp(-\alpha|x - \mu|)$$

1. Write a code generating double exponential distribution $DE(0,1)$ from $U(0,1)$ by using inverse CDF method. Explain how you obtained that code step by step. Generate 10000 random numbers from this distribution, plot the histogram and comment whether the result looks reasonable.
2. Use Acceptance/rejection method with $DE(0,1)$ as majorizing density to generate $N(0,1)$ variables. Explain step by step how this was done. How did you choose constant c in this method? Generate 2000 random numbers $N(0,1)$ using your code and plot the histogram. Compute the average rejection rate R in the acceptance/rejection procedure. What is the expected rejection rate ER and how close is it to R ? Generate 2000 numbers from $N(0,1)$ using standard *rnorm* procedure, plot the histogram and compare the obtained two histograms.

Submission procedure

Assume that X is the current lab number.

If you are neither speaker nor opponent for this lab,

- Submit your report using *Lab X* item in the *Submissions* folder before the deadline.
- Make sure that you or some of your group members submits the group report using *Lab X group report* in the *Submissions* folder before the deadline

If you are a speaker for this lab,

- Submit your report using *Lab X* item in the *Submissions* folder before the deadline.
- Make sure that you or some of your group members does the following before the deadline:
 - submits the group report using *Lab X group report* in the *Submissions* folder before the deadline
 - Goes to Study room *Speakers X* → *Documents* and opens file *Password X.txt*. Then the student should put your group report into ZIP file *Lab X.zip* and protect it with a password you found in *Password X.txt*
 - Uploads the file to *Collaborative workspace* folder

If you are opponent for this lab,

- Submit your report using *Lab X* item in the *Submissions* folder before the deadline.

- Make sure that you or some of your group members submits the group report using *Lab X group report* in the *Submissions* folder before the deadline
- After the deadline for the lab has passed, go to Collaborative workspace folder and download *Lab X.zip*. Open the PDF in this ZIP file by using the password available in *Course Documents* → *Password X.txt*, read it carefully and **prepare at least two questions/comments/improvement suggestions** in order to put them at the seminar (i.e. at least two questions per opponent)